

R E P O R T R E S U M E S

ED 017 958

72

CG 001 797

A BATTERY OF TESTS ON GENERAL EDUCATIONAL DEVELOPMENT FOR
POST-ELEMENTARY SCHOOLS. VOL. II.

BY- BENTWICH, J. AND OTHERS

HENRIETTA SZOLD INSTITUTE, JERUSALEM (ISRAEL)

REPORT NUMBER HSI-RR-NO-112

PUB DATE

MAR 67

RUTH BRESSLER CENTER FOR RESEARCH IN EDUCATION

CONTRACT OEC-4-21-D11

EDRS PRICE MF-\$0.25 HC-\$1.64 39P.

DESCRIPTORS- *TEST CONSTRUCTION, *HEBREW, *ACHIEVEMENT TESTS,
*APTITUDE TESTS, GRADE 9, GRADE 10, TEST RELIABILITY, TEST
VALIDITY,

THIS BATTERY OF TESTS WAS DEVELOPED TO MEASURE THE
GENERAL EDUCATIONAL DEVELOPMENT OF STUDENTS IN SECONDARY
SCHOOLS IN ISRAEL. BY DEVELOPING AN INSTRUMENT THAT MEASURED
GENERALIZED SKILLS AND ABILITY TO USE INFORMATION RATHER THAN
A STRICT ACHIEVEMENT TEST, IT WAS HOPED THAT THE INSTRUMENT
COULD BE USED FOR BOTH GUIDANCE PURPOSES AND THE IMPROVEMENT
OF EVALUATION METHODS. THE BATTERY CONSISTS OF SIX TESTS-(1)
MATHEMATICS, (2) SCIENCE, (3) READING
COMPREHENSION-LITERATURE, (4) READING COMPREHENSION-SOCIAL
STUDIES, (5) SOCIAL STUDIES-GENERAL INFORMATION, AND (6)
ENGLISH LANGUAGE. RELIABILITY DATA INCLUDED INTERCORRELATIONS
OF THE SUBTEST AND CORRELATIONS OF THE SUBTEST WITH THE TOTAL
TEST SCORE. THE RELIABILITY OF THE TOTAL SCORE EXCEEDED .90
AND MET ACCEPTABLE STANDARDS. CONTEST VALIDITY WAS
ESTABLISHED BY THE JUDGMENT OF TEACHERS, SCHOOLS SUPERVISORS,
AND EXPERTS ON THE SUBJECT MATTER, WHILE CORRELATIONS BETWEEN
TEST SCORE AND GRADE-POINT AVERAGE WERE USED TO DETERMINE
CONCURRENT VALIDITY. HOWEVER, VALIDITY WITH RESPECT TO
PREDICTION OF ACADEMIC ACHIEVEMENT, SUCCESS IN SCHOOL, AND
EFFECTIVENESS OF THE TESTS FOR TREATMENT AND PLACEMENT
PURPOSES STILL REQUIRES A FOLLOW-UP STUDY. THE NORMS WERE
BASED ON A SAMPLE OF CLASSES, GRADES NINE AND 10, IN
ACCREDITED ACADEMIC HIGH SCHOOLS. A MORE DETAILED DISCUSSION
OF THE TESTS IS GIVEN IN HEBREW. (CG)

ED017958

THE HENRIETTA SZOLD INSTITUTE
NATIONAL INSTITUTE FOR RESEARCH
IN THE BEHAVIORAL SCIENCES

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE
PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS
STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION
POSITION OR POLICY.

No BR no.

PA72

Contract no.

OE 4-21-64

A BATTERY OF TESTS
ON GENERAL EDUCATIONAL DEVELOPMENT
FOR POST-ELEMENTARY SCHOOLS

BY

J. BENTWICH, J. LEVIN, D. ORMIAN

RESEARCH REPORT No. 112

PUBLICATION No. 456

CG 001 797

MARCH 1967

THE HENRIETTA SZOLD INSTITUTE

National Institute for Research in the Behavioral Sciences

Project on Construction of a Set of Instruments
in order to Assess Achievements, Aptitudes,
and Attitudes of Students in Secondary
Schools in Israel

b y

J. BENTWICH, J. LEVIN, D. ORMIAN

This project was made possible by a grant from
The United States Government , Department of Education
Project No. OE-4-21-011

TABLE OF CONTENTS

	Page
I. DEVELOPMENT OF THE PROJECT	2
II. DESCRIPTION OF THE BATTERY	3
III. VALIDITY	6
IV. DISTRIBUTION AND NORMS	7
V. REFERENCES	10

LIST OF TABLES

1. Intercorrelations, means and standard deviations of the subtests for grade nine	4
2. Intercorrelations, means and standard deviations of the subtests for grade ten	5
3. Subtest - Total Score Correlations	6

LIST OF FIGURES

A. Distribution of Scores (average of six subtests in grades nine and ten)	8
--	---

A BATTERY OF TESTS OF EDUCATIONAL DEVELOPMENT

Project on Construction of a Set of Instruments
in order to Assess Achievements, Aptitudes, and
Attitudes of Students in Secondary Schools
in Israel

This is a summary of the technical report on the project. A detailed discussion and description of the tests is given in Hebrew.

I. Development of the Project

The initial plan of the project was prompted by a series of research publications on current examinations and methods of assessment in high schools. These studies emphasized the inadequacies of the evaluation methods with respect to reliability, distribution of scores and standardization. Surveys of school failures and drop-outs, published at about the same time, showed serious bottlenecks in the secondary school system (see references on p. 10).

There is a close relation between these issues, since success and failure in school, among other things, depend on the evaluation methods employed by the teachers, and on the availability of guidance aided by proper instruments.

Adequate assessment methods are also essential for adaptation of curriculum to the level of students and their rate of progress. Since the structure of assessment instruments should conform to these purported objectives, it is important to keep one main objective in the foreground. There is consensus of opinion among educational experts emphasizing guidance as the most urgent task.

Since standardized intelligence tests are already available in Hebrew, it was decided to focus on educational attainment. Owing to the diversity of curricula in the lower grades of the high school, it seemed preferable to construct a test battery to measure general educational development rather than strict achievement tests. Also, from the guidance point of view, the general level of a student is far more important than specific items of information.

The well known batteries I T E D (Iowa Tests of Educational Development) and S T E P (Sequential Tests of Educational Progress) served as a model for the present battery:

"...THE TESTS ARE DESIGNED TO MEASURE MUCH MORE THE GENERALIZED SKILLS. THEY ARE INTENDED TO MEASURE THE PUPIL'S ABILITY TO DO CRITICAL THINKING IN THE BROAD AREAS DESIGNATED. THEY ARE CONCERNED NOT SO MUCH WITH WHAT THE PUPIL HAS LEARNED, IN THE SENSE OF SPECIFIC INFORMATION, BUT RATHER WITH HOW WELL HE CAN USE WHATEVER HE HAS LEARNED IN ACQUIRING, INTERPRETING AND EVALUATING NEW IDEAS, IN RELATING NEW IDEAS TO OLD, AND IN APPLYING BROAD CONCEPTS AND GENERALIZATIONS TO NEW SITUATIONS ON THE SOLUTION OF NEW PROBLEMS. THESE ARE SOME OF THE OUTCOMES NOT ONLY OF AN EFFECTIVE COURSE OF FORMAL SCHOOL INSTRUCTION, BUT ALSO OF ANY OTHER GENUINELY EDUCATIONAL EXPERIENCE, WHETHER FORMAL OR INFORMAL, DIRECT OR INCIDENTAL, IN-SCHOOL OR OUT-OF-SCHOOL..." (from Manual for Teachers and Counselors, I T E D).

The construction and standardization of the battery followed the usual stages of pilot study, pretest for item analysis and final normative study. The reliabilities, intercorrelations and validity reported here are based on the data of the normative sample. The sampling procedure is described in Chapter IV.

II. Description of the Battery and Intercorrelations

The battery comprises six subtests: 1) Mathematics 2) Science (Physics, Chemistry, Biology) 3) Reading Comprehension - Literature 4) Reading Comprehension-Social Studies 5) Social Studies - General Information 6) English Language.

Mathematics and English are achievement, rather than educational development tests. Their inclusion in the battery is due to their central key position in the high school curriculum.

A test applied for guidance purposes ought to satisfy high standards of reliability. Owing to their brevity, reliabilities of subtests are not high enough for construction of profiles of individual students. Higher reliabilities could be obtained even for short subtests by selection of items maximizing subtest homogeneity but this policy would miss the aim of the battery to cover broad aspects of educational level. However, the reliability of the total score exceeds .90 and meets acceptable standards.

The intercorrelations, means and standard deviations of the subtests are given in Tables 1 and 2. The entries in the diagonal are the split-half reliabilities corrected by the Spearman-Brown Formula. The entries above the diagonal are the raw correlations. Below the diagonal are the correlations corrected for attenuation.

TABLE 1
Intercorrelations, means and standard deviations
of the subtests for grade nine
(N = 370)

Subtest	1 Mathe- matics	2 Science	3 Reading Comprehension Literature	4 Reading Comprehension Social Studies	5 Social Studies- General Information	6 English
1 Mathematics	(.75)	.52	.27	.30	.45	.32
2 Science	.73	(.68)	.29	.37	.47	.29
3 Reading Comprehension- Literature	.42	.47	(.57)	.47	.37	.35
4 Reading Comprehension- Social Studies	.48	.62	.85	(.53)	.41	.37
5 Social Studies- General Information	.63	.68	.59	.68	(.69)	.28
6 English	.40	.38	.50	.55	.36	(.85)
Mean	13.6	17.5	13.6	13.6	12.6	33.4
Standard Deviation	4.0	4.3	2.8	3.1	4.3	7.8

TABLE 2

Intercorrelations, means and standard deviations
of the subtests for grade ten
(N = 370)

Subtest	1 Mathematics	2 Science	3 Reading Comprehension- Literature	4 Reading Comprehension- Social Studies	5 Social Studies- General Information	6 English
1 Mathematics	(.72)	.49	.31	.36	.37	.25
2 Science	.68	(.71)	.29	.37	.32	.38
3 Reading Comprehension- Literature	.46	.43	(.65)	.51	.36	.35
4 Reading Comprehension- Social Studies	.55	.57	.82	(.60)	.36	.40
5 Social Studies- General Information	.54	.47	.55	.57	(.66)	.21
6 English	.32	.48	.47	.56	.28	(.87)
Mean	15.3	19.0	14.0	14.3	13.4	39.9
Standard Deviation	4.0	4.8	3.2	3.0	4.0	8.0

The pattern of the correlations conforms to those commonly obtained. The highest correlations are between the reading comprehension tests (literature and social studies). This is due to similarity of function and content. But a thorough investigation of the composition of the battery would require a factor analytic study.

The contribution of each subtest to the total score is measured by its correlation. The correlations of the subtests with the total score are given in Table 3.

TABLE 3
Subtest - Total score correlations

Subtest	<u>Grade 9</u>	<u>Grade 10</u>
1 Mathematics	.68	.65
2 Science	.69	.71
3 Literature	.60	.63
4 Social Studies - Reading	.65	.68
5 Social Studies - General Information	.69	.59
6 English	.75	.76

The range of the relative weights i.e. the ratios of the correlations is between 1.0 and 1.3. The tests may be considered as equally weighted - the largest weight is of the English test which is only slightly larger than the others (1.3 in both grades).

III. Validity

The content validity of the battery was established by judgement of teachers, school supervisors, and experts on the subject matter of each subtest respectively.

Concurrent validity with respect to academic achievement is based on correlations between the battery score and point-grade averages of 21 classes. The median correlation is .50. The interquartile range is .32. These values underestimate the validity because of the restriction of range in each class. This becomes evident when we note the large difference in scores among schools. However, computation of the correlation between the battery scores and school grades for the whole sample would involve standardization of school grades on one consistent scale.

The validity with respect to prediction of academic achievement, success and failure in school, dropping-out and effectiveness of the tests for adaptive treatment and placement purposes, requires a follow-up study.

Some possible applications of the battery are related to adjustment problems. Significant discrepancies between test scores and school grades may indicate adjustment difficulties related to motivation, interests, personality and teacher-student relations. This issue also requires further investigations. An intensive validation and follow-up study is now being planned and will be undertaken soon.

IV. Distribution and Norms

The distribution of scores of the battery in grades nine and ten are given in figure A.

The norms are based on a sample of classes, grade nine and grade ten, in accredited academic high schools. The sampling procedure was based on a stratification of schools by type of settlement and level of school (based on the results of matriculation examinations). The total number of cases was 370 from each grade. The Sample covered 21 classes. The data were collected during the final trimester of 1966.

The means of the distributions are 16.9 and 18.8 in grades nine and ten respectively. The standard deviations are 3.3 and 3.9 in grades nine and ten respectively.

The large overlap of the distributions shows that individual differences far exceed the differences between grade levels. Similar results can be seen by comparison of class means. For example, on the Social Science General Information subtest the means of the median class of grade nine and grade ten are 12.3 and 14.3 respectively, the difference being 2.0, while the range of the class means of grade nine and of grade ten are 7.5 and 5.3 respectively. Similar results occur in the other subtests.

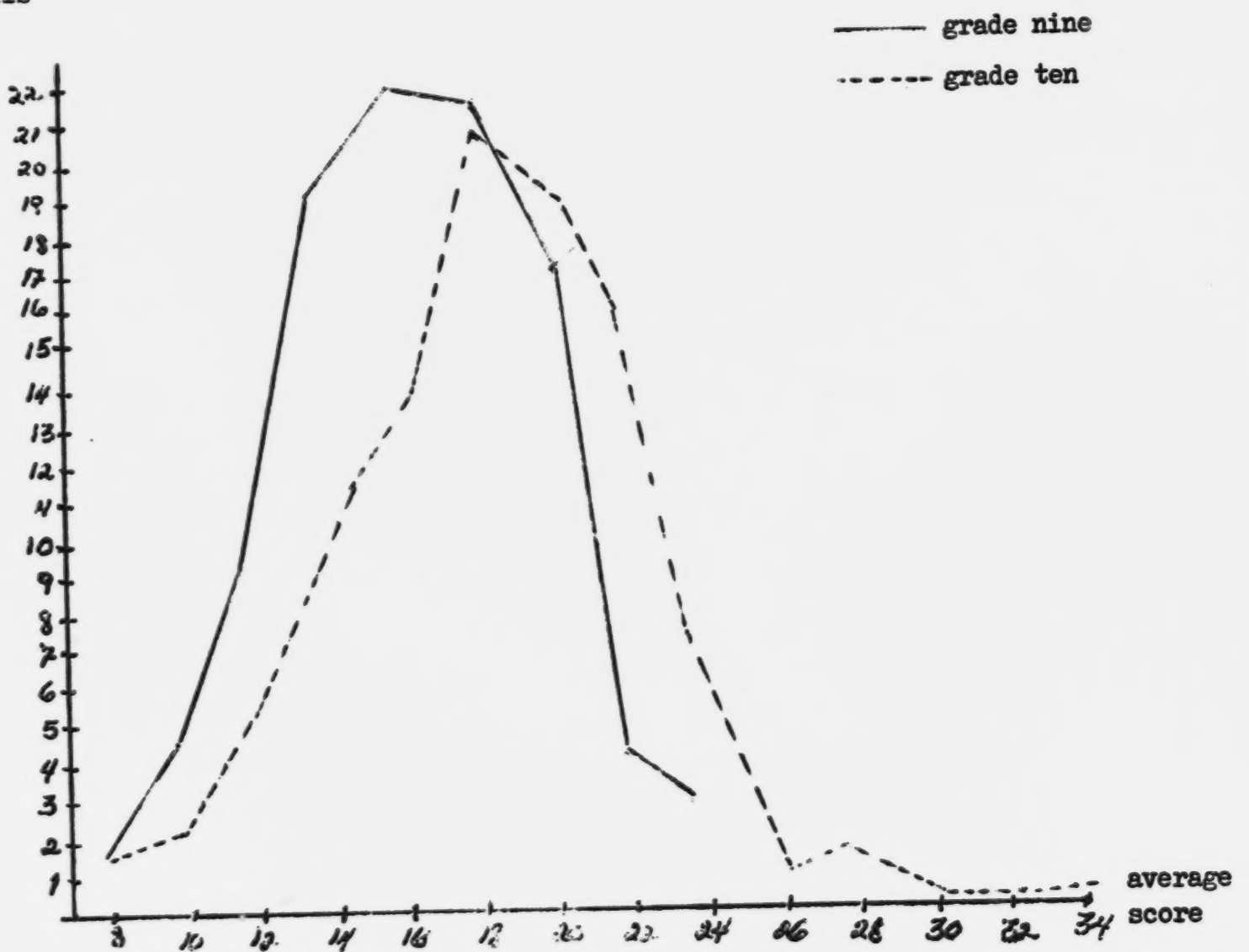
This is an expected outcome especially for a battery measuring general educational development, for attainment here depends on the background of the pupils, methods of instruction, extent of individual treatment, encouragement of extracurricular activities, and the selective policies of various schools rather than on progress during one year of high school. However, the available data are not sufficient for a thorough analysis of the factors involved.

For guidance purposes it is necessary to extend the norms in order to cover vocational schools and agricultural schools. This will be done in conjunction with the forthcoming validation and follow-up study.

Figure A

Distribution of scores (average of six subtests) in grade nine and in grade ten, in percentages

percentage
of pupils



Mean of average score: grade nine - 16.9, grade ten - 18.8 .
Standard deviation of score: grade nine - 3.3, grade ten - 3.9 .

In addition to the description and explanation referred above, the Hebrew part of this report includes the following tables:

- Table I - Reliability of the subtests and of the total battery
- Table III - Ratios between the total score correlations
- Table VI - Distribution of the class means of each subtest - grade nine
- Table VII - Distribution of the class means of each subtest - grade ten
- Table VIII - Number of points necessary to add or subtract to scores of pupils participating in 5 subtests
- Table IX - Norms and percentile scores of the mean score

V. REFERENCES

1. Bentwich, J.: Examinations and evaluation in secondary schools. Megamot, 1963.
2. Eson, M. E., Bernstein, R. J.: The curriculum and problems of evaluation. Megamot, 1958.
3. Guttman, E. L., Ortar, G.: Examination of validity of entrance examinations for candidates of the School of Medicine. The Institute for Social and Applied Research, Jerusalem.
4. Muhsam, H. V., Klinov-Mallul, R., Chanoch, G.: The supply of professional manpower from Israel. Educational Sources, Jerusalem, 1959.
5. Orr, L.: The reliability of the Israeli Bagrut examinations. Megamot, 1963.
6. Pilliner, A. E. G.: The Measurement of Educational Achievement in Israel, 1961.
7. Smilansky, M.: Outline for reform in secondary education. Megamot, 1961.

מכון הנריאטה סאלד
המכון הארצי למחקר במדעי ההתנהגות

מערכת מבחנים לבתי"ס על-יסודיים

ע"י

יוסף בנטואיץ, יוסף לויין ודליה אורמیان

**ביצוע מחקר זה התאפשר הודות לעזרת משרד
החנוך בממשלת ארצות הברית**

• זכרונות על ידי
משה שניידר
המחבר
המחבר

תוכן הענינים

עמוד

1	א. תיאור מערכת המבחנים
4	ב. נתונים סטטיסטיים: המדגם, מהימנות, תקפות, אינטרקורלציות, זמן
12	ג. השימושים במערכת המבחנים
16	ד. התפלגות הציון במערכת, נורמות
20	ה. הוראות להערכת מערכת המבחנים

רשימת הלוחות

5	I. מהימנות תת-המבחנים וכלל המערכת
6	II. המיתאמים בין הציונים בתת-המבחנים לבין הציון בכלל המערכת
7	III. היחס בין המיתאמים
10	IV. אינטרקורלציות בין תת-המבחנים - כיחה ס'
10	V. אינטרקורלציות בין תת-המבחנים - כיחה י'
14	VI. התפלגות הכיחות לפי הציון הממוצע של הכיחה בכל תת-מבחן - כיחה ס'
14	VII. התפלגות הכיחות לפי הציון הממוצע של הכיחה בכל תת-מבחן - כיחה י'
17	VIII. מספר הנקודות שיש להוסיף או לחסיר לציוני זלמידים שנבחנו ב-5 תת-מבחנים
18	IX. נורמות של הציון הממוצע

רשימת דיאגרמות

17	1. התפלגות באחוזים של הציון הממוצע במערכת המבחנים
----	---

א. תיאור מערכת המבחנים

מערכת המבחנים נועדה לשמש כמכשיר עזר בהכוון ובייעוץ לתלמידי הכיתות ט' ו-י' של בתי"ס על-יסודיים.*

המערכת כוודת את רמת ההתפתחות החינוכית הכללית של התלמיד, המותנה הן בגורמים של הישג בלימודים והן ברמת המישכל. היא עומדת איפוא במקום ביניים בין ציוני ביה"ס - המודדים בעיקרו של דבר את ההישג בלימודים, לבין מבחני אינטלגנציה, המודדים את רמת המישכל.

בפיתוח המערכת עמדה לנגד עינינו מערכת המבחנים ITED שנבנתה במכללת איובה

בארה"ב. להלן העקרונות המנחים ששמשו כבסיס לבניית ITED.

"...המבחנים מיועדים למדידת מיומנות כללית. מטרתם למדוד את יכולתו של התלמיד לחשוב

חשיבה בקורטית בתחומים שונים. הם קשורים אמנם ללימודים של ביה"ס אולם עיקרו של קשר זה אינו במדידת ידע ספציפי, כי אם במדידת היכולת ליישם את הידע שנרכש, היכולת לפרש ולהעריך רעיונות חדשים, לייחס ולקשר רעיונות חדשים וישנים, ליישם מושגים והכללות למצבים חדשים ולפתרון בעיות. יכולת זו היא תוצאה הן של ההוראה בביה"ס במובן הפורמאלי, והן של הנסיון הנרכש, במובן החינוכי, במסגרת פורמאלית או בלתי פורמאלית, בדרך ישירה או בעקיפין, בביה"ס או מחוצה לו."

(על פי: ITED: The Iowa tests of educational development: a manual for teachers and counselors).

כצעד ראשון, הוכן מכשיר לבחי-ספר עיוניים.

אומדן רמתו של התלמיד מבחינת התפתחות חינוכית כללית עשוי לסייע בייעוץ ובהכוון מהבחינות הבאות:

ראשית - קביעת התאמתו של התלמיד לכיתה ולביה"ס. מכשירי מדידה מדויקים חשובים במיוחד לגבי תלמידים הדורשים טיפול, אם מפני קשיים בלימודים, אם מפני בעיות משמעת, חוסר-סיפוק מצד התלמיד, וכיו"ב.

* בהמשך המחקר נאסוף נתונים גם על בתי"ס בלתי-עיוניים.

במקרים רבים אפשר למנוע נשירת תלמידים ממערכת ההינוך ע"י הפנייתם למסלול אחר, למגמה אחרת או לבי"ס אחר המתאים יותר להתפתחותם הכללית. יש להעיר אמנם כי ההתפתחות הכללית של התלמיד איננה הגורם היחיד שיש לקחת בחשבון. הפנייה לבי"ס מקצועי למשל, חייבת להתבסס גם על זיהוי כושר ספציפי להצלחה במקצוע הנלמד וכן על קיום נטיות ותכונות אישיות מתאימות.

שנית, פער גדול בין הציונים בלימודים לבין הרמה הכללית הנמדדת במערכת זו, רומז על בעיות הסתגלות של התלמיד למסגרת החינוכית. אפשרות זו מחייבת תשומת לב מיוחדת, ומצריכה לעתים הפנייה של התלמיד לתחנות לייעוץ.

בנוסף לעימוש במבחנים לשם הערכת התלמיד היחיד, אפשר להשתמש בהם גם להערכת כיתות ולהשוואה ביניהן (ר' פירוש בפרק ג').

מערכת המבחנים כוללת 6 תת-מבחנים במקצועות: מתמטיקה, מדעי הטבע, הבנת הנקרא - קטעי ספרות, הבנת הנקרא - קטעים במדעי החברה, שאלות במדעי החברה, אנגלית.

יש הבדלים בין המבחנים במידת הקשר שלהם לתכנית הלימודים:

התלות בתכנית הלימודים גבוהה יותר יחסית במבחנים במקצועות הריאליים (מתמטיקה ומדעי הטבע) ובמבחן באנגלית. לעומת זאת, המבחנים במקצועות ההומניסטיים (ספרות, 2 מבחנים במדעי החברה) קשורים יחסית פחות לתכנית הלימודים, ומודגשים בהם יותר הגורמים של ידע כללי ושל הבנת מושגי יסוד. ברור שההדגשה על קשר לתכנית לימודים מכאן או על ידע כללי מאידך, איננה שווה בכל הפריטים של מבחן.

המבחנים הכלולים במערכת:

(1) מתמטיקה - 24 שאלות בגיאומטריה ובאלגברה.

(2) מדעי הטבע - 33 שאלות בפיסיקה, בכימיה ובביולוגיה. השאלות קשורות אמנם לחומר

הלימודים בביה"ס, אולם התעניינות אישית של התלמיד בתחומים אלה יכולה להרחיב את מספר השאלות שידע להשיב עליהן.

ברוב בתי"ס לא לומדים את כל שלושת המקצועות הכלולים במבחן זה. השתדלנו איפוא

לאזן את המבחן ע"י מתן ייצוג מתאים לכ"א מהתחומים. כיוון שהנורמות מבוססות על תוצאות

שהושגו ע"י תלמידים שברובם אינם לומדים את כל המקצועות, באה עובדה זו לידי ביטוי

בנורמות. תלמידים שלא למדו כימיה יוכלו להגיע איפוא לציונים דומים לשל תלמידים שאינם לומדים פיסיקה, ע"י הישגים טובים בעיקר בשאלות הנוגעות לתחום אותו למדו.

(3) ספרות* - המבחן בודק הבנת קטעים הלקוחים מהספרות. כלולים במבחן 6

קטעים עם שאלות אחדות לגבי כל קטע. סה"כ - 20 שאלות.

(4) חברה - הבנת קטעי קריאה. נבדקת הבנת קטעים ממדעי החברה. המבחן כולל 5

קטעים עם 22 שאלות.

(5) חברה - שאלות. 22 שאלות בהיסטוריה ובחולדות ישראל, בכלכלה, באזרחות

ובגיאוגרפיה. חלקן קשור בחומר הלימודים וחלקן מותנה בעיקר בהשכלה הכללית של התלמיד.

ההערה שהובאה בהקשר למבחן במדעי הטבע יפה גם כאן: ברוב בחה"ס אין לומדים את כל התחומים הכלולים במבחן, אך הנורמות לוקחות זאת בחשבון כיוון שהן מבוססות על מדגם ארצי.

(6) אנגלית - 55 שאלות הכוללות אוצר מלים, דקדוק והבנת קטעי קריאה.

הפריטים במבחנים ערוכים בדרך כלל לפי סדר קושי.

* קטעי הספרות הם משל:

ספר האגדה; ש.ה. ברגמן; ל. גלודברג; א. קובנר; ג. שופמן; י. שנברג.

ב. נתונים סטטיסטיים:

1. המדגם
2. מהימנות
3. תקפות
4. אינטרקורלציות
5. זמן

1. המדגם

המדגם שמשמש בסיס לחישוב הנורמות הוא מדגם ארצי של כיתות ט' ו-י' בבח"ס על-יסודיים עיוניים מוכרים.

המדגם הוא מדגם שכבות, כשבתוך כל שכבה הוצא מדגם מקרי. הריבוד נעשה לפי: (1) צורת-הישוב של המקום בו נמצא בית הספר (3 הערים הגדולות, ערים אחרות; יתר צורות הישוב); (2) הרמה בלימודים של ביה"ס. מאחר שאין מדד אחיד ומקובל לקביעת הרמה, בחרנו לצורך זה את המודד של אחוז המצליחים בבחינות הבגרות בשנה שקדמה למחקר (חשכ"ה). החלוקה היא לשתי רמות.

המדגם מקיף 370 תלמידים בכיתה ט' (11 כיתות) ו-370 תלמידים בכיתה י' (13 כיתות). בהמשך המחקר בכוונתנו לחשב נורמות גם עבור בח"ס על-יסודיים לא-עיוניים. המבחנים הועברו בשליש האחרון של שנת הלימודים תשכ"ו.

2. מהימנות

מהימנות המבחן מודדת את מידת הדיוק של התוצאות. כלומר: באיזו מידה היו משיגים התלמידים את אותו הציון במבחן מקביל. כיוון שאין בידינו מבחן מקביל, חלקנו את הפריטים בכל מבחן לשני חתאים - פריטים זוגיים ופריטים בלתי-זוגיים, וחישבנו את המיתאם בין הציונים של שני חצאים אלה. כ"כ חישבנו את המהימנות הכוללת של המערכת.

התוצאות נחונות בלוח I.

לוח I. מהימנויות תת-המבחנים וכלל המערכת*

מבחן	כיתה ט'	כיתה י'
מתמטיקה	.75	.72
מדעי הטבע	.68	.71
ספרות - קטעים	.57	.65
חברה - קטעים	.53	.60
חברה - שאלות	.69	.66
אנגלית	.85	.87
כלל המערכת	.90	.91

* מתוקן לפי נוסחת ספירמן-בראון.

המהימנות הגבוהה ביותר היא של המבחן באנגלית (.87 - .85). הדבר נובע כנראה מכך שזהו המבחן הארוך ביותר, וכידוע תלויה המהימנות באורך המבחן. המהימנויות של המבחנים במקצועות הריאליים (.75 - .68) גבוהות מאלה של המקצועות ההומניים (.69 - .53). המהימנויות של המבחנים הבודדים אינן עומדות בדרישות המקובלות ממבחן שנועד לאבחנה אינדיבידואלית. מהימנות של מבחן בודד תלויה הרבה יותר - מאשר מהימנות כלל המערכת - בגורמים מקריים (למשל, חומר הלמודים הספציפי שנלמד בכיתה, סעויות מקריות). רוב המבחנים הם קצרים למדי ומהימנותו של מבחן עולה עם אורכו. אפשר להבין קשר זה אם נשים לב שמשקלם של כשלון או הצלחה מקריים בשאלה בודדת, גדול יותר במבחן קצר מאשר במבחן ארוך. יש להעיר שהאורך האפקטיבי של המבחנים (כלומר, תחום הציונים במבחן), קטן אף ממספר השאלות שהם מכילים, כיוון שהשאלות הראשונות בכל מבחן קלות מאד וכולס פותרים אותן. נהוג לכלול שאלות כאלה במבחן לשם עידוד הנבחנים.

לאור האמור לעיל החלטנו להתייחס אל ההישג של התלמיד בכלל המערכת. צירוף

המבחנים נותן לנו מערכת בת 176 שאלות. במערכת כזו מאזנות התנודות המקריות זו את זו, ומכאן המהימנויות הגבוהות יותר - 90. בכיתה ט' ו-91. בכיתה י', שהן בהתאם לסטנדרד המקובל.

יש להעיר כי ניתן היה להגיע למהימנויות גבוהות יותר אילו הרכבנו מבחן הומוגני יותר, המקיף תחומים מצומצמים יותר. אלא שמבחן מסוג כזה היה מחטיא את המטרה של מדידה גלובאלית. נראה לנו איפוא כי עדיף להרכיב מבחן המקיף תחומים רבים יותר, אף אם הוא מהימן פחות.

ההצלחה במערכת המבחנים תימדד ע"י הציון הממוצע במערכת. בהתאם לכך, הנורמות שחושבו הן נורמות של הציון הממוצע במערכת. לפי גישה זו יש למבחנים הארוכים השפעה רבה יותר על הציון מאשר למבחנים הקצרים.

שיטה מדויקת יותר היתה מחייבת חישוב ציונים עפ"י ממוצע משוקלל, אולם שיטת החישוב היתה מסתבכת. מאחר שמתברר שההבדלים במשקלים של המבחנים במערכת קטנים למדי, לא נראה לנו כדאי לסבך את שיטת החישוב.

משקלו של מבחן במערכת נמדד ע"י מקדם המיתאם בין הציון במבחן לבין הציון במערכת כולה.

II. המיתאמים בין הציונים בתת-המבחנים לבין הציון בכלל המערכת

<u>מבחן</u>	<u>כיתה ט'</u>	<u>כיתה י'</u>
מחמטקה	.68	.65
מדעי הטבע	.69	.71
ספרות - קטעים	.60	.63
חברה - קטעים	.65	.68
חברה - שאלות	.69	.59
אנגלית	.75	.76

היחס בין המשקלים נתון בלוח III (למיתאם הקטן ביותר בכל דרגת כיחה נתנו ערך של 1).

לוח III. היחס בין המיתאמים

מבחן	כיתה ט'	כיתה י'
מחמטקה	1.1	1.1
מדעי הטבע	1.2	1.2
ספרות - קטעים	<u>1.0</u>	1.1
חברה - קטעים	1.1	1.2
חברה - שאלות	1.2	<u>1.0</u>
אנגלית	1.3	1.3

כפי שמראה לוח III למבחנים משקל דומה. המבחן באנגלית הוא בעל משקל גבוה

קצת יותר, אולם ההבדל איננו רב, ואיננו מצדיק את הסיבוך הכרוך בשינוי שיטת השקלול.

3. תקפות

תקפותו של מבחן מוגדרת כמדת ההלימה בין המבחן לבין התופעה אותה מיועד למדוד. שיטות ההערכה של התקפות מורכבות למדי משום שבדרך כלל אין קריטריון אחד-משמעי למדידת התקפות. למעשה אפשר לקבוע את תקפותו של מבחן רק על סמך מחקר ממושך ועל סמך סידרה של השוואות והערכות, אשר מצטרפות לחמונה מקיפה של יחסים ומיתאמים בין המבחן לבין נתונים מעשיים ותאורטיים.

נוהגים להבחין בין סוגים אחדים של תקפות:

(1) תקפות תוכן - תקפות המשתמעת מעיון בתוכנם של פריטי המבחן. בדרך כלל קובעים

את תקפות התוכן של מבחן על פי הערכת מומחים. למשל, מבחני הישגים הצמודים לתכנית לימודים יש להעריך לפי חוות-דעת של מחנכים, הקובעים אם אמנם משקפים הפריטים את התכנית ואת הציפיות לגבי ההישגים.

הפריטים של מערכת מבחנים זו חוברו ע"י מורים. כמו כן נחקלה חוות-דעתם של בעלי מקצוע בתחומים המתאימים הן על הפריטים והן על מידת האיזון בין התחומים השונים. במובן זה עומדת המערכת בדרישות של תקפות התוכן.

(2) תקפות מבנית - תקפות המושתתת על מערכת יחסים ומיתאמים במסגרת של מודל תיאורטי. למעשה מעריכים תקפות זו לפי המשמעות והסדירות של לוח האינטרקורלציות של מערכת המבחנים או עפ"י ניתוח גורמים. התקפות המבנית נידונה בסעיף 3 בפרק זה. מערכת האינטרקורלציות הולמת את תוכנם של המבחנים ומתאימה למצופה.

(3) תקפות מקבילה (concurrent) - המיתאם בין המבחן לבין קריטריון הנמדד בו בזמן שנערך המבחן. למשל, מיתאם בין מבחן הישגים לבין הערכות מורים. חישבנו את המיתאם שבין הציון הממוצע במערכת המבחנים לבין הציון הממוצע של המורה בסוף שנת הלימודים. ציון המורה הוא ממוצע המקצועות הבאים:

1. מחמסיקה.

2. מדעי הטבע - ממוצע הציונים בפיסיקה, בכימיה ובביולוגיה (בכיתות הלומדות רק חלק ממקצועות אלה: ממוצע של המקצועות הנלמדים).
3. ספרות - ממוצע הציונים בתנ"ך ובספרות.
4. מדעי החברה - ממוצע הציונים בהיסטוריה, בגיאוגרפיה ובאזרחות (בכיתות הלומדות רק חלק ממקצועות אלה: ממוצע של המקצועות הנלמדים).
5. אנגלית.

חישוב המיתאם נערך בכל כיתה בנפרד. המיתאם החציוני הוא 0.50, והתחום הבין-רביעוני של המיתאמים הוא 0.32.

מבחינה עקרונית, המיתאם לגבי האוכלוסיה כולה גבוה הרבה יותר: תחום הציונים במערכת המבחנים בחוץ כל כיתה קטן למדי, בהשוואה לתחום הציונים במערכת באוכלוסיה כולה. צמצום התחום מקטין בהרבה את גובה המיתאם.

עם זאת אין הצדקה לחישוב המיתאם לגבי כלל האוכלוסיה, כיוון שסולם ציוני המורים שונה מכיתה לכיתה (יש מורים המחמירים בציונים, ויש המקלים. ציון 6 בכיתה אחת אינו מקביל לציון 6 בכיתה אחרת). חישוב המיתאם לגבי כלל האוכלוסיה היה מצריך כיול של ציוני המורים.

(4) תקפות חיזוי - יכולתו של מבחן לחזות תופעות ותהליכים לטווח ארוך.

למשל, יכולתו של מבחן לחזות נשירה מביה"ס, הצלחה בעבודה וכו'.

אין בידינו עדיין נתונים בנוגע ליכולתה של המערכת לחזות גורמים כמו הצלחה בלימודים, נשירה מביה"ס וכד'. בדיקת תקפות זו כרוכה במעקב ממושך, והנתונים ייאספו בהדרגה בשיתוף עם בתי הספר.

4. אינטרקורלציות

המיתאמים בין המבחנים נתונים בלוחות IV ו-V. המספרים באלכסון בסוגריים הן המהימנויות (מתוקנות לפי נוסחת ספירמן-בראון). מעל לאלכסון נתונים המיתאמים הגולמיים; מתחת לאלכסון נתונים מיתאמים מתוקנים לפי נוסחה המבטלת את השפעת הבדלי המהימנויות, כלומר: נותנת את המיתאמים שהיו מתקבלים אילו המהימנות היתה מגיעה לאחד.

לוח IV . אינטרקורלציות בין תת-המבחנים - כיתה ט'

תת-מבחן	1 מחמטיקה	2 מדעי הטבע	3 ספרות קטעים	4 חברה קטעים	5 חברה שאלות	6 אנגלית
1. מחמטיקה	(.75)	.52	.27	.30	.45	.32
2. מדעי הטבע	.73	(.68)	.29	.37	.47	.29
3. ספרות - קטעים	.42	.47	(.57)	.47	.37	.35
4. חברה - קטעים	.48	.62	.85	(.53)	.41	.37
5. חברה - שאלות	.63	.68	.59	.68	(.69)	.28
6. אנגלית	.40	.38	.50	.55	.36	(.85)
ציון ממוצע	13.60	17.50	13.61	13.55	12.58	33.43
סטית תקן	4.01	4.32	2.82	3.10	4.35	7.85

לוח V . אינטרקורלציות בין תת-המבחנים - כיתה י'

תת-מבחן	1 מחמטיקה	2 מדעי הטבע	3 ספרות קטעים	4 חברה קטעים	5 חברה שאלות	6 אנגלית
1. מחמטיקה	(.72)	.49	.31	.36	.37	.25
2. מדעי הטבע	.68	(.71)	.29	.37	.32	.38
3. ספרות - קטעים	.46	.43	(.65)	.51	.36	.35
4. חברה - קטעים	.55	.57	.82	(.60)	.36	.40
5. חברה - שאלות	.54	.47	.55	.57	(.66)	.21
6. אנגלית	.32	.48	.47	.56	.28	(.87)
ציון ממוצע	15.32	18.96	14.03	14.28	13.42	39.88
סטית תקן	3.95	4.85	3.16	3.04	4.05	7.95

המיתאמים הגבוהים ביותר שנתקבלו (ר' מיתאמים מתחת לאלכסון) הם מיתאמים

בין:

1. חברה - קטעים לספרות - קטעים (.85 - .82). שני אלה הם מבחנים של הבנת

קטעי קריאה; גם מבחינת התוכן מבחנים אלה דומים למדי.

2. מחמטיקה ומדעי הטבע (.73 - .68) - שני המבחנים במקצועות הריאליים

במערכת זו.

יתר המיתאמים נמוכים יותר. הנמוכים ביותר הם המיתאמים עם אנגלית.

5. זמן

בפרק ה' מובא פירוט הזמן שהוקצב לכל מבחן. זמן זה נקבע עפ"י ניסוי מוקדם

שנערך בכיתות אחדות. מקובל לקבוע כאופטימום פרק זמן המאפשר ל-85% מהתלמידים לסיים

את המבחן; לפי זה נקבעו הזמנים של כל תת-מבחן, כזמן בו סיימו את המבחן 85%

מהתלמידים.

המבחן תלוי איפוא רק במידה קטנה במהירות. לא נהוג להקציב זמן שבו יגמרו הכל

את המבחן, כיוון שתמיד מצויים תלמידים שמשתהים יתר על המידה. תוספת זה מעבר למה

שנקבע כמעט ואיננה משנה את דירוג הציונים של התלמידים ורק משבשת את סדרי העברת

המבחן. יחד עם זאת על הבוחן להקפיד על הזמן מאחר וכל הבדיקות הסטאטיסטיות והנורמות

מחושבות לפי הקצבת הזמן המפורטת בהוראות.

ג. השימושים במערכת המבחנים

מערכת המבחנים נועדה, כאמור, לשמש מכשיר עזר בייעוץ ובהכוון של תלמידים בכיתות ט' ו-י' בבח"ס על-יסודיים. הנורמות הנתונות בלוח IX מאפשרות לקבוע את רמתו של התלמיד בהשוואה לסטנדרד ארצי בבחי-ספר עיוניים.

קביעת הרמה חשובה במיוחד לגבי תלמידים הזקוקים לייעוץ ולהכוונה. למשל, לגבי תלמידים שלגביהם קיימים ספקות אם יוכלו לעבור לכיתה גבוהה יותר או לגבי תלמידים שעל סף הנשירה. ציונים והערכות מורים המכירים את התלמיד במשך שנה או שנתיים משמשים כמובן גורם מכריע בשיקולים אלה, אולם מכשיר מדידה מדויק, המכויל על פי נורמה ארצית, עשוי לסייע במידה רבה להבהרת התמונה של התלמיד הבעייתי.

במקרים רבים אפשר למנוע נשירתו של תלמיד ממערכת החינוך ע"י הפנייתו למסגרת המתאימה לרמתו. בהקשר זה יש להדגיש כי מערכת מבחנים זו איננה שלמה ואיננה מספיקה כדי לייעץ לתלמיד לעבור למסגרת של בי"ס בלתי עיוני. ייעוץ כזה צריך לקחת בחשבון גם כשרים ספציפיים של התלמיד, נטיות ותכונות אישיות שלו. התוצאה במבחנים היא רק גורם אחד בין הגורמים שיש לקחת בחשבון. במקרים בעייתיים יש להעביר את המקרה לידי פסיכולוג-יועץ או לתחנה להדרכה מקצועית.

מבחינה פסיכולוגית מעניינים במיוחד מקרים של פער גדול בין ציוני ביה"ס לבין התוצאה במערכת המבחנים. במקרים אחדים אפשר לייחס פער זה לסטיוה מקרית, שמקורה באי דיוק במדידה. תוצאות של מבחן מסוים עלולות להיות מושפעות לרעה מהפרעות בריכוז או ממתיחות יתר. מצד שני הן עשויות להיות מושפעות לטובה מידע מקרי של פריטים מסוימים או מניחוש מוצלח. הוא הדין גם לגבי הערכות המורים אשר אינן בעלות דיוק מוחלט.

אולם בדרך כלל מעיד הפרש ניכר בין הרמה במערכת המבחנים לבין ציוני המורים על בעיות הסתגלות של התלמיד למסגרת ביה"ס - בין אם אלה בעיות מוטיבאציה ובין אם אלה בעיות של קונפליקטים נפשיים. תלמידים מסוג זה זקוקים לחשומת לב מיוחדת מצד הפסיכולוג-היועץ של ביה"ס.

נוסף לשימוש במערכת לגבי מקרים בעייתיים, היא מהווה מכשיר הערכה לדירוג תלמידי

הכיתה.

אפשר לדרג את התלמידים לפי הציון הגלמי שלהם במערכת; אולם קל יותר להבין את משמעות הציונים אם מתרגמים אותם לנורמות (ר' לוח IX) מאחר והנורמות מתרגמות את הציונים לסולם בהחאס לסטנדרדים ארציים של ציוני המורים.

ההערכה המדויקת ביותר למקומו של התלמיד בכיתתו מבוססת על הנורמה של הציון הממוצע שלו במערכת ועל ציון המורים. שיטת חישוב הנורמה מפורטת בפרק ד'.

כ"כ אפשר לדרג את הכיתה על בסיס של מדגם ארצי של כיתות:

מתוך הציון הממוצע של התלמידים אפשר לחשב את הציון הממוצע של הכיתה בכל אחד מתת-המבחנים. לפי הנתונים שבלוחות VI ו-VII אפשר לאתר לאיזה קבוצת כיתות שייכת הכיתה.

למשל, אם הציון הממוצע של כיתה ט' מסוימת במבחן במתמטיקה היה 13.3, פירושו שהכיתה שייכת במבחן זה לרבע השלישי (מלמטה) של הכיתות. כיתה י' בעלת ציון ממוצע של 13.1 בספרות - שייכת לרבע השני (מלמטה) של הכיתות.

כאמור, הציון במבחן אחד איננו מהימן במידה מספקת כדי לקבוע על פיו אבחנה אינדיוידואלית. אולם הממוצע של כיתה יציב יותר, מאחר והסטיות המקריות של טעות המדידה מבטלות זו את זו, ולכן אפשר לקבוע את מקומה של כיתה בין כלל הכיתות בכל מבחן לחוד.

לוח VI . התפלגות הכיתות לפי הציון הממוצע של הכיתה בכל תת-מבחן - כיתה ט'

קבוצת הכיתות	תת-מבחן	מתמטיקה	מדעי הטבע	ספרות - קטעים	חברה - קטעים	חברה - שאלות	אנגלית	ציון ממוצע במערכת
רבע ראשון (הנמוך ביותר)	עד 12.3	עד 15.7	עד 12.7	עד 12.7	עד 12.7	עד 10.3	עד 29.8	עד 16.4
רבע שני	12.4-13.1	15.8-17.0	12.8-13.7	12.8-13.9	10.4-12.2	29.9-33.6	16.5-17.1	
רבע שלישי	13.2-14.9	17.1-19.6	13.8-14.3	14.0-14.2	12.3-14.0	33.7-35.6	17.2-18.3	
רבע רביעי (הגבוה ביותר)	15.0 ומעלה	19.7 ומעלה	14.4 ומעלה	14.3 ומעלה	14.1 ומעלה	35.7 ומעלה	18.4 ומעלה	

לוח VII . התפלגות הכיתות לפי הציון הממוצע של הכיתה בכל תת-מבחן - כיתה י'

קבוצת הכיתות	תת-מבחן	מתמטיקה	מדעי הטבע	ספרות - קטעים	חברה - קטעים	חברה - שאלות	אנגלית	ציון ממוצע
רבע ראשון (הנמוך ביותר)	עד 13.2	עד 16.6	עד 13.0	עד 13.0	עד 12.7	עד 35.5	עד 18.1	
רבע שני	13.3-16.0	16.7-17.8	13.1-14.1	13.1-13.8	12.8-14.2	35.6-38.8	18.2-19.2	
רבע שלישי	16.1-16.3	17.9-21.0	14.2-14.3	13.9-14.9	14.3-14.4	38.9-44.3	19.3-21.2	
רבע רביעי (הגבוה ביותר)	16.4 ומעלה	21.1 ומעלה	14.4 ומעלה	15.0 ומעלה	14.5 ומעלה	44.4 ומעלה	21.3 ומעלה	

ההבדלים בין בתה"ס גדולים במידה ניכרת מההבדלים בין דרגות כיתה באותם בת"ס.
למשל, במבחן חברה-שאלות: הציון של כיתה ט' החציונית הוא 12.3, הציון של כיתה י'
החציונית הוא 14.3, וההפרש איפוא - 2.0; לעומת זה תחום הציונים בחברה-שאלות בתוך
כיתות ט' הוא בין 7.5 נקודות (הכיתה הנמוכה ביותר קבלה 9.2 נקודות, והגבוהה ביותר -
16.7 נקודות. נתונים אלה אינם מופיעים בלוח VI), ובחור כיתות י' בין 5.3 נקודות.
(מ-10.6 ועד 15.9). תופעה דומה נתקבלה גם ביחיד המבחנים.

יחד עם זאת, מספר הכיתות במדגם (11 כיתות ט' ו-13 כיתות י') אינו גדול במידה
מספקת כדי לנתח תוצאה זו לפרטיה.

ד. התפלגות הציון במערכת; נורמות

כאמור, חושבו הנורמות על בסיס של מדגם ארצי של כיתות: 11 כיתות ט', (370 תלמידים) ו-13 כיתות י' (370 תלמידים), של בח"ס על-יסודיים עיוניים מוכרים. הנורמות הן של סוף שנת הלימודים תשכ"ו.

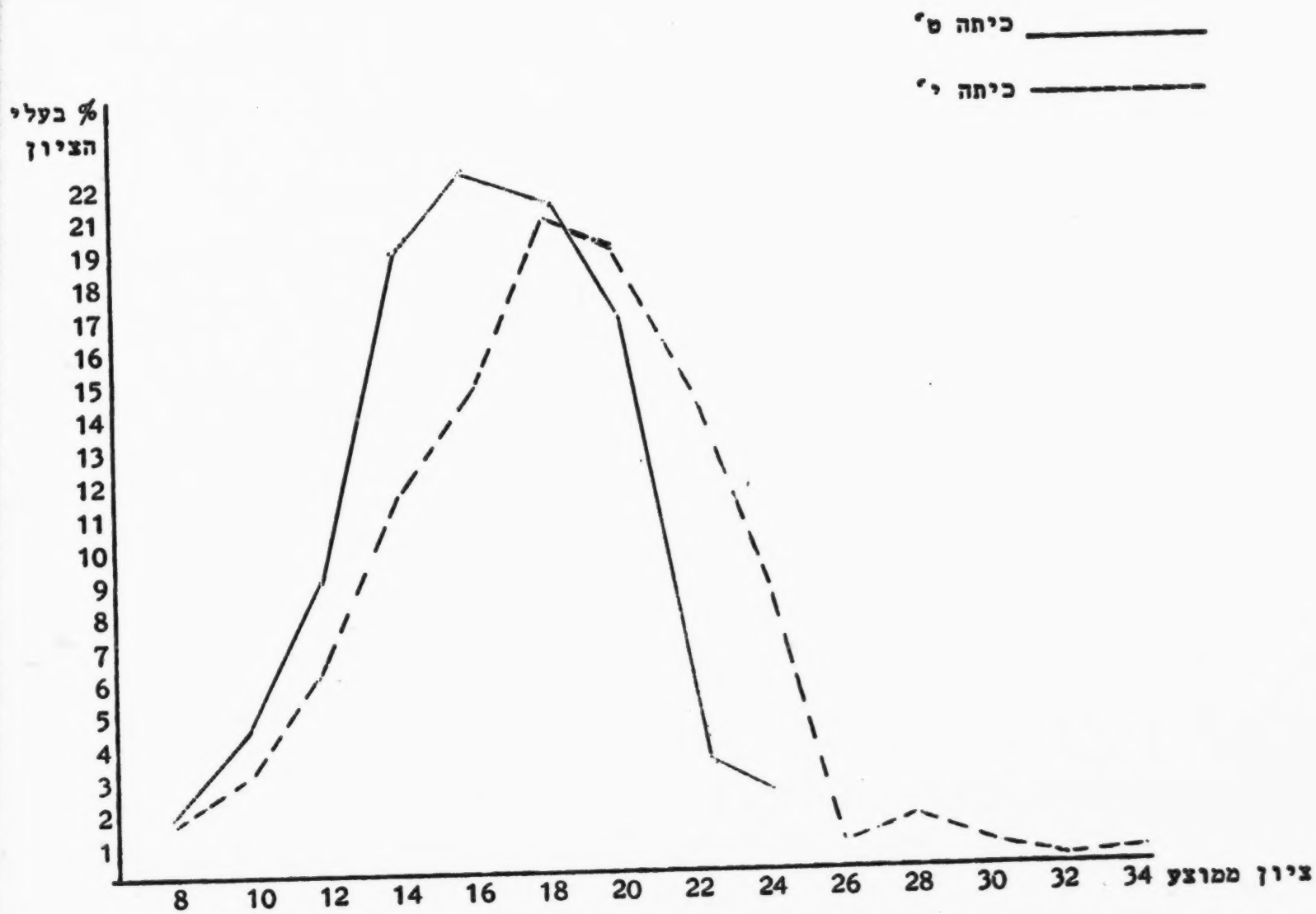
הנורמות מתייחסות לציון הממוצע של התלמיד בכלל המערכת. הממוצע מחושב ע"י סה"כ הציונים של התלמיד ב-6 המבחנים, חלקי 6 (מעוגל למספרים שלמים). דיאגרמה 1 מתארת את ההתפלגות באחוזים של הציון הממוצע בכיתות ט' ו-י'. הציון הממוצע הוא 16.9 בכיתה ט' ו-18.8 בכיתה י', וסטיית החזק 3.3 בכיתה ט' ו-3.9 בכיתה י'. הדיאגרמה מראה כי תלמידי כיתות י' עולים אמנם בדרך כלל על תלמידי כיתות ט', אולם קיימת חפיפה ניכרת בין ההתפלגויות, וההבדלים האינדיווידואליים בין תלמידים באותה דרגת כיתה גדולים מההבדלים בין ההתפלגויות של שתי דרגות הכיתה.

יתכן שהסיבה לכך היא אופיו של המבחן, הבודק רמת התפתחות כללית ואינו צמוד לתכנית הלימודים. ההבדלים הנובעים משנת לימודים אחת אינם משתקפים כאן באופן בולט, בהשוואה להבדלים שמקורם ברקע התלמידים, בחינוכם הקודם ובהבדלים ברמתם הכללית של בתי הספר. בדרכי הוראה וחינוך - סיפוח דרכי חשיבה, עידוד קריאה רחבה, ארגון חיי חברה וחוגים, טיפול בתלמידים החלשי וכיו"ב. אך כאמור הנתונים שבידינו אינם מספיקים לניתוח מעמיק יותר של התוצאה.

במקרה שהתלמיד החסיר מבחן אחד או כשהמבחן נפסל מסיבות טכניות או מקריות - יש לחשב את הממוצע של חמשת המבחנים ולהוסיף או להחסיר לתוצאה את הגדלים הבאים (את העיגול למספרים שלמים יש לעשות לאחר התיקון):

דיאגרמה 1:

התפלגות הציון הממוצע במערכת המבחנים, באחוזים



ממוצע הציונים הממוצעים: ט' - 16.9 י' - 18.8

סטיית תקן של הציונים הממוצעים: ט' - 3.3 י' - 3.9

לוח VIII , מספר הנקודות שיש להוסיף או להחסיר לציוני
תלמידים שנבחנו ב-5 תת-מבחנים

המבחן החסר	הפעולה	כיתה ס'	כיתה י'
מחמטקה	להחסיר	.8	.8
מדעי הטבע	-	-	-
ספרות - קטעים	להחסיר	.8	1.1
חברה - קטעים	להחסיר	.8	1.0
חברה - שאלות	להחסיר	1.0	1.2
אנגלית	להוסיף	3.2	4.1

מטרת תיקונים אלה היא לפצות את ההשפעה של ההבדלים במוצעי המבחנים. למשל, תלמיד שהחסיר מבחן באנגלית יפסיד יותר נקודות מתלמיד שהחסיר מבחן בספרות. התיקונים נותנים איפוא את ההפרשים בין הממוצע (של כלל האוכלוסיה) ב-6 מבחנים לבין הממוצעים (של כלל האוכלוסיה) ב-5 מבחנים.

יש לזכור שלוח הנורמות מבוסס על ממוצעי 6 מבחנים. השימוש בלוח לגבי ממוצע 5 מבחנים גורם אי דיוק מסוים. גם המהימנות של ממוצע של 5 מבחנים יורדת במידת מה. במקרה שחסרים 2 מבחנים או יותר - אין להסתמך על המערכת לאבחנה אינדיוידואלית.

לוח IX . נורמות של הציון הממוצע

כיתה י'		כיתה ט'		ציון ממוצע במערכת המבחנים
% מצטבר של תלמידים שהשיגו ציון זה או נמוך ממנו	סטנדרד מקביל בציוני מורה	% מצטבר של תלמידים שהשיגו ציון זה או נמוך ממנו	סטנדרד מקביל בציוני מורה	
100	89 ומעלה			31 ומעלה
99	84			30
99	84			29
99	84			28
98	82			27
98	82			26
98	82	100	87 ומעלה	25
95	79	99	84	24
90	76	97	81	23
82	71	96	81	22
76	69	94	79	21
65	66	87	75	20
57	65	77	71	19
46	63	66	68	18
35	61	56	65	17
26	59	44	62	16
21	58	34	60	15
14	57	23	58	14
9	56	15	55	13
6	54	10	53	12
4	53	6	50	11
2	49	3	48	10
1	46	1	45	9
0	עד 46 (כולל)	1	עד 45 (כולל)	עד 8 (כולל)

בלוח הנורמות שלעיל מוצג הציון הממוצע במערכת המבחנים (טור ראשון) בהקבלה ל:

(1) האחוז המצטבר של תלמידים שהשיגו ציון זה או נמוך ממנו (כיתה ט' - טור

שלישי. כיתה י' - טור חמישי).

למשל, תלמיד הכיתה ט' שהשיג ציון ממוצע של 20 נקודות במערכת, הוא מעל 87%

תלמיד כיתה ט'. במלים אחרות: מתוך 100 תלמידים בדרגת כיתה ט' עולים עליו בהישג

במערכת המבחנים 13 תלמידים, ונמוכים ממנו - (או שווים לו) 87 תלמידים.

(2) ציון ממוצע של המורים בסוף שנת הלימודים (כיתה ט' - טור שני. כיתה י' -

טור רביעי). כמוזכר קודם, חישבנו עבור כל תלמיד שנכלל במדגם ציון ממוצע של המורים

על סמך ציוני סוף השנה (על דרך חישוב הציון - ר' פרק ב' - סעיף 2).

על סמך התפלגות ציוני המורים במדגמנו, ערכנו הקבלה בינם לבין הציון הממוצע

במערכת. הקבלה זו מאפשרת לנו לבטא את הציון במערכת המבחנים במונחים המוכרים והרווחים

של ציון המורים.

למשל, ציון ממוצע במערכת המבחנים של 15 נקודות לתלמיד בכיתה י', מקביל לציון

ממוצע שניתן ע"י מורים של קצת פחות מ"מספיק" (58). 15 נקודות במערכת ו-58 נקודות

בביה"ס, הם הנקודה המפרידה בין 21% התלמידים הנמוכים לבין 79% התלמידים הגבוהים.

כיוון שמדובר בממוצע ציון מורה, יש לשים לב שתחום הציונים איננו מ-40 עד 100,

אלא מ-45 ועד 89. מכאן שציון המעיד על כשלון יהיה ציון המקביל לציון מורה של 44

(ולא 40), ואילו ציון המעיד על הצטיינות יהיה זה המקביל ל-89 (בכיתה י', או 87

בכיתה ט'). ההסבר ההגיוני לתופעה סטטיסטית זו הוא שעל מנת לקבל ציון ממוצע 4 על

התלמיד להיכשל בכל המקצועות והיפוכו לגבי התלמיד המצטיין. תופעה כזו נדירה כ"כ עד כי

לא נמצאו תלמידים כאלה במידגם. עיון בטור האחוזים יעזור למורה לפרש את התוצאה מבחינה

אינטואיטיבית יותר מאשר עיון בטור הציונים.

יש להעיר שציוני המורה אינם כוללים מקצועות הקשורים לידע או לכושר ספציפי,

כגון - ציור, מוסיקה, התעמלות וכד', כיוון שגם מערכת המבחנים איננה מתייחסת

לאספקטים אלה. כשרים מסוג זה נבדקים במבחנים מיוחדים המיועדים לכך.

ה. הוראות להעברת מערכת המבחנים

1. מערכת המבחנים מתאימה להעברה הן באופן קבוצתי - לכל הכיתה יחד, והן באופן אינדיווידואלי.

2. המבחן יועבר ע"י מורה הכיתה. רצוי שבכיתה יהיה נוכח עוד מישהו. יש להושיב את הנבחנים עם רוח מספיק ביניהם.

יש להעיר בהקשר זה כי על הבוחן להקפיד ולדאוג מראש לתנאים טכניים נוחים (סדר ישיבה מתאים, לוודא שלכל התלמידים יש כלי כתיבה, וכד'), כיוון שזהו מבחן עם הגבלת זמן, יש לדאוג לעניינים אלה מראש, ולמנוע הוצרות בעיות בעת המבחן עצמו.

3. המבחנים יועברו בשני מועדים, עם הפסקה של ימים אחדים בין מועד למועד.

4. לפני תחילת המבחן יקדים הבוחן את דברי ההסבר הבאים. מטרת ההסבר היא למנוע

מתח הנובע מההרגשה שהמבחן גורלי, אך ליצור יחד עם זאת גישה רצינית כלפיו.

דברי ההסבר:

אנו עומדים להעביר בכיתה 6 מבחנים - שלשה היום ושלשה ביום... המבחנים בודקים הן ידיעות במקצועות שונים והן השכלה כללית. בין השאלות יש גם שאלות קשות על נושאים שלא למדתם ושלא קראתם עליהם. הסיבה היא שהמבחן מיועד גם לכיתות בעלות תכנית לימודים שונה מזו שלכם. יחד עם זה תמצאו מספיק שאלות קלות יותר, שתוכלו לענות עליהן.

בכמה מבחנים לא כולם יצליחו לסיים את המבחן. במקרה שתסיימו את המבחן לפני תום הזמן העומד לרשותכם, רצוי שתעברו עליו ותבדקו שנית את תשובותיכם.

סדר העבודה יהיה כזה:

אחרי שאחלק את המבחנים נקרא יחד את ההוראות. לאחר מכן תוכלו לשאול שאלות. כשאודיע על כך - תהפכו את הדפים וכל הכיתה תתחיל יחד את המבחן. אם מתעוררות שאלות בזמן המבחן - הצביעו, ואגש אליכם. אענה רק על שאלות סידוריות, ולא על שאלות הנוגעות לתוכן המבחן. כשאודיע לכם להפסיק יניחו כולם את העטים ויפסיקו - גם אם הם באמצע שאלה. אחרי כל מבחן נעשה הפסקה קצרה.

המבחן הראשון הוא מבחן במדעי החברה. את ההוראות נקרא יחד לאחר שאחלק אותו.

5. הבוחן יחלק לתלמידים את השאלון הראשון עם תשובות. התלמידים ימלאו את שמם

בראש התשובות. הבוחן יקרא בקול את ההוראות למבחן הראשון.

יש לדאוג שכל התלמידים יתחילו יחד את המבחן ויפסיקו יחד.

6. לפני המבחנים הבאים יש להזכיר בקיצור את ההוראות העיקריות. בתחילת המועד השני

של הבחינה - כדאי לחזור שוב על ההסבר.

7. סדר המבחנים והזמן המוקצב לכל אחד הוא:

מועד א'

1. חברה - שאלות - 15 דקות

2. מתמטיקה - 50 דקות

3. ספרות - 30 דקות

סה"כ 1.35 שעה. עם הסבר והפסקות: כשעתיים.

את המבחן במתמטיקה כדאי לתת מיד לאחר המבחן בחברה - שאלות. בין מתמטיקה לספרות -

לעשות הפסקה של כ-10 דקות. במבחן למתימטיקה: לחלק גם נירות סיוטה.

מועד ב'

4. חברה - קטעים - 30 דקות

5. מדעי הטבע - 30 דקות

6. אנגלית - 30 דקות

סה"כ 1.30 שעה. עם הסבר והפסקות: פחות משעתיים.

כדאי לתת את המבחן במדעי הטבע מיד לאחר חברה - קטעים. לפני אנגלית: לעשות הפסקה

של כ-10 דקות. במבחן באנגלית אין תשובון - התשובות נכתבות בגוף המבחן.

8. בדיקת המבחנים תיעשה לפי מפתח התשובות הנכונות הרצ"ב. הבדיקה תיעשה על גבי

התשובונים. שאלה שקבלה שתי תשובות נחשבת כשגויה.

9. הציון של תלמיד בכל מבחן הוא מספר התשובות הנכונות שלו באותו מבחן.

הציון הממוצע של התלמיד מחושב ע"י סה"כ הציונים בששת המבחנים - מחולק ב-6.

לוח הנורמות (לוח IX) מדרג את התלמיד באוכלוסית התלמידים בדרגת כיחתו, ומתרגם את

הישגיו לסולם מקביל לציוני-מורה.

מפתח התשובות: מבחנים 1-5

מספר השאלה	חזק-מבחן	ה ת ש ו ב ה				ה נ כ ו נ ה	
		חברה שאלות	מחמסיקה	ספרות	חברה - קטעים	מדעי הטבע	
2		ב	ה	ב	ב	א	
3		ג	ג	א	א	ד	
4		ד	א	ג	ד	ד	
5		ב	ג	ד	א	ד	
6		ג	ד	ג	א	ד	
7		ד	ב	א	ג	ג	
8		ב	ג	ג	ב	ב	
9		א	א	ב	ב	א	
10		ג	ב	ד	ד	ב	
11		ב	ב	א	ג	ב	
12		ב	ב	ד	ב	ב	
13		ג	ד	ד	א	ב	
14		ג	ג	ב	ד	ג	
15		ג	ג	ב	ג	ג	
16		ב	ה	ב	ד	ב	
17		א	א	ב	ג	ג	
18		ב	ב	א	ב	ג	
19		א	א	ד	א	א	
20		א	ה	ב	ב	ג	
21		ב	ב	א	ד	ד	
22		ג	ג		ג	א	
23		ג	ד		ג	א	
24			ד			ב	
25			ג			ד	
26						ב	
27						א	
28						ג	
29						א	
30						א	
31						א	
32						ג	
33						א	

מפתח תשובות: מבחן באנגלית

<u>מספר השאלה</u>	<u>חשובה נכונה</u>	<u>מספר השאלה</u>	<u>חשובה נכונה</u>
1	Have you got	21	ד
2	Does John like	22	ב
3	Did he give Tom	23	ב
4	Does Mary help	24	ג
5	on	25	א
6	than	26	ד
7	than	27	ב
8	at	28	ד
9	D	29	ג
10	B	30	א
11	C	31	א
12	A	32	ג
13	D	33	ד
14	B	34	א
15	D	35	ג
16	D	36	נכון
17	water	37	נכון
18	how much	38	לא נכון
19	you to	39	נכון
20	writing	40	לא נכון
		41	לא נכון
		42	לא נכון
		43	נכון
		44	לא נכון
		45	נכון
		46	לא נכון
		47	לא נכון
		48	נכון
		49	נכון
		50	לא נכון
		51	נכון
		52	לא נכון
		53	לא נכון
		54	לא נכון
		55	לא נכון